A simple model for the control of cell-type proportions in multicellular development(Mathematical Topics in Biology)

AUTHOR(S):

INOUYE[INOUE], Kei

CITATION:

INOUYE[INOUE], Kei. A simple model for the control of cell-type proportions in multicellular development(Mathematical Topics in Biology). 数理解析研究所講究録 1991, 762: 53-65

ISSUE DATE:

1991-07

数理解析研究所講究録
第762巻 1991年 53-65

A simple model for the control of cell-type proportions in multicellular development

INOUYE Kei 井上 敬
Department of Botany, Faculty of Science, Kyoto University
Sakyo-ku, Kyoto 606 Japan

Abstract

A simple conceptual model is proposed for the generation of multiple cell-types from an initially homogeneous population of cells. In the model, the state of each cell is defined by its gene expression pattern and physiological parameters, the former of which being discrete corresponding to different cell types, whereas the latter being governed by differential equations representing the physicochemical laws. Using a simplified version of the model, requirements for the coexistence of different cell types within a cell population, and factors influencing their proportions, are studied.

Introduction

Since the pioneering work by Turing (1952), a number of attempts have been made to simulate the formation of biological patterns by the use of mathematical models (for reviews, Meinhardt, 1982; Murray, 1989; Nagorcka, 1989). Most models proposed so far concentrate on generating the spatial distribution of the imaginary substance which is supposed to induce a definite cell differentiation or the formation of specific structures ("morphogen" after Turing) whereas the ability of proportion regulation seems to have been treated as a subordinate property to be possessed by the pattern generating mechanisms.

There are cases, however, in which the cells differentiate in a definite proportion but without any particular spatial pattern. Even in the case where a clear spatial pattern arises, there are examples in which the cells differentiate first without taking any particular spatial arrangement but sort themselves out afterwards to generate a coherent pattern. The control of cell-type propor-

tions and the formation of spatial patterns are therefore conceptually separable, and it will be important, especially in theoretical studies, to make this distinction clear. In addition, despite the important contributions of the theoretical models, they have been often criticized on the ground of biological reality. In fact, the formulation of the model, *i.e.* the specifications of the variables and their interactions, are often arbitrary.

To extend the usefulness of mathematical models in the study of developmental biology, it is desired to construct a conceptual framework for theoretical models on accepted biological grounds. In this study I attempted to do this by reducing the problems to as simple a model as possible. It is not intended in the present paper to give a detailed description of the model but only to put forward some basic ideas on which mathematical theories of developmental phenomena might be constructed. Here we will focus on the control of cell-type proportions. Formation of spatial patterns by different cell-types will be investigated elsewhere.

Main questions

The elemental processes constituting the control of cell-type proportions are:

(a) A specific set of cells expresses a specific set of genes, which defines a specific cell-type.

(b) Different cell-types coexist in the same developmental system.

(c) Proportions of different cell-types are controlled.

(d) When a part of the system is removed, the remaining part restores itself to the normal proportion.

We will define by these statements the following terms, respectively; (a) cell differentiation, (b) diversification, or "division of labour", (c) cell-type proportioning, and (d) proportion regulation. We will examine what are required for each of these phenomena to take place using a simple model system consisting of idealized cells.

The model

Division of cell state into two components

We can postulate with little loss of generality that all the cells of the developmental system under consideration are genetically identical. However, individual cells may, despite their having an identical genetic information, take different "states". The change in the state of the cells is primarily determined by the cell-autonomous dynamics, but will also be influenced by the other cells of the same system.

Since our primary interest is in cell differentiation, it may seem legitimate to define the state of a cell by the genes that are being transcribed in that cell. However, the cellular activities are mostly carried out by a huge number of molecules and ions which constitute the physical entity of the cell. Considering that cells, in ordinary development, influence each other not by direct DNA-DNA interactions but via changes in the physiological parameters, we are led to incorporate physiological variables explicitly in the model.

The state of a cell is defined in the model by specifying (i) the genes that are being transcribed and (ii) the values of the physiological variables. For the convenience's sake, we will call the former component "*gene state* (G)" and the latter "*physiological state* (P)". Corresponding to the division of the cell state into two components, the dynamics of the cell state is divided into two parts: one (which we designate by u) governing the change in the former (which change may be called "*developmental change*") and the other (v) the change in the latter (which may be called "*physiological change*").

How the cell state changes

Since the dynamics of the physiological state is after all chemical reactions occurring in a very complex situation, its change should be basically smooth, being described, in principle, by a set of partial differential equations. On the contrary, the change in the gene state is discrete, being a succession of

ONs and OFFs of the genes. The dynamics of the physiological change is strongly restricted by the geometry of the cell and the enzymes active at each instant. The latter element (and partly the former element also) of the restriction is the effect of the genes that have been expressed. On the other hand, transcription of any gene is believed to be controlled by cellular components other than DNA itself, such as DNA-binding proteins. In other words, the dynamics of the gene state and that of the physiological state are imposing strong constraints on each other:

$$G_i(t + dt) = u(G_i(t); P(t)) \tag{1a}$$

$$\frac{dP(t)}{dt} = v(P(t); G_i(t)). \tag{1b}$$

A change in the gene state is, be it an ON of new genes or OFF of the genes that have been expressed, induced when the physiological state moves, in due course of its innate dynamics, into the domain where the prior gene state is no longer stable. Any change of the gene state in turn must alter the dynamics of physiological state and deviate the course of the physiological change from the one it would otherwise take, and such a deviation could be substantial if the gene that was switched on or off encoded a key enzyme of some reaction network. Fig.1 shows changes in the state of a hypothetical cell having only three genes, A, B, C, and one physiological variable P. In this imaginary and highly simplified situation, P first increases slowly, with the dynamics defined by genes B and C which are being expressed, to reach a point P_1 where gene A is turned on. The dynamics has now changed due to the effect of gene A so that P starts to increase more rapidly. When P reaches a point P_2, gene B is switched off and P now changes with a new dynamics determined by genes A and C. This example illustrates how the gene state and physiological state interact each other within a cell.

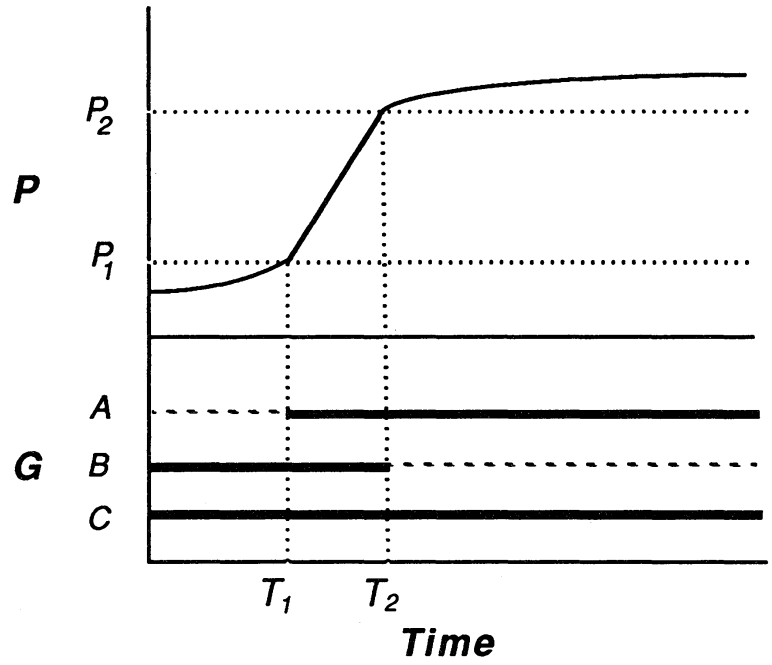

Fig.1. A schematic diagram illustrating how the cell state changes. The state of the hypothetical cell is represented by the gene state (G) and the physiological state (P). In the model the former is determined by the expression of the genes in question (A and B) and a group of common genes (C), and its physiological change is assumed to be described by a single parameter P. The genes that are being expressed are indicated by thick lines.

Cell interaction and division of labour

Interaction between the cells is, as pointed out earlier, mediated by changes in the physiological state of these cells. For the system comprising N cells, the dynamics of cell state is given by

$$G_i(t+dt) = u(G_i(t);\, P_i(t)) \tag{2a}$$

$$\frac{dP_i(t)}{dt} = v(P_i(t);\, G_i(t), P_1, P_2, \cdots, P_N). \tag{2b}$$

$$i = 1, 2, \cdots, N$$

In reality, only a limited number of the components of P_i will be involved in the interaction of cells.

To illustrate how cell interaction affects the cell state, consider a system comprising but two cells (cells *1* and *2*) of the kind described earlier. Here,

however, we assume that only gene C is expressed by the cells initially, and that gene A is turned on when P reaches P_1 whereas B becomes ON if P decreases to P_2. If there is no interaction between the cells, the time-course of the state change will be identical (or nearly so if we allow a limited variation between cells) to each other and to the one for a solitary cell. With a moderate interaction, there may arise some modulation such as a delay or acceleration of the expression of A. Strong interaction, however, may give rise to such a situation that the physiological state of, say, cell *1* reaches the critical point P_1 slightly earlier than cell *2* (Fig.2). If the change in the physiological state of cell *1* due to the expression of gene A is such that it prevents the physiological state of cell *2* from attaining P_1 by, for instance, forcing P of cell *2* to decrease, cell *2* will not express gene A and eventually gene B may be

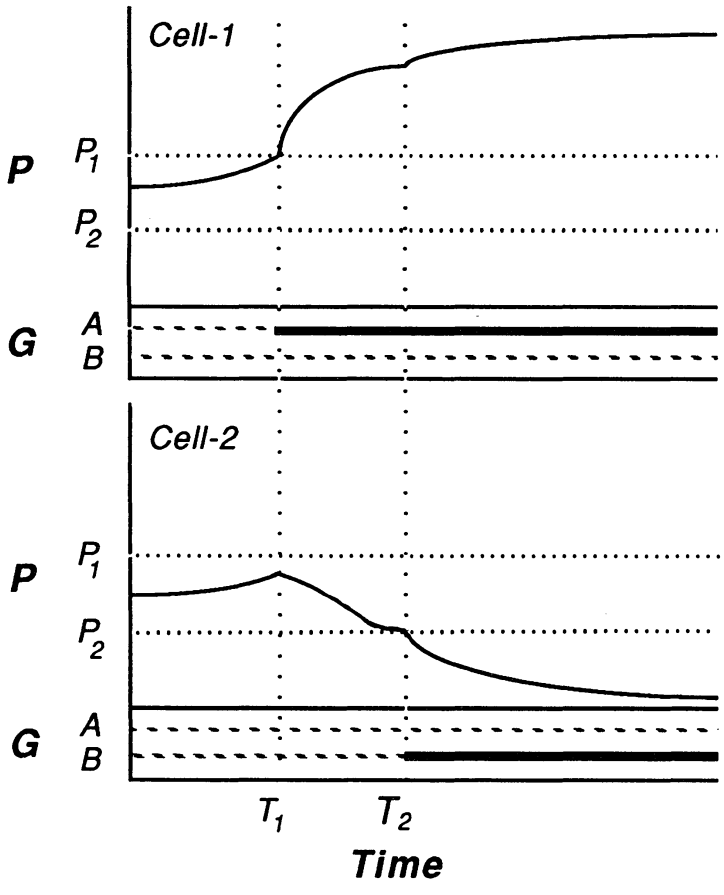

Fig.2. An example showing how cell interaction may give rise to division of labour.

switched on, which is a situation representing the simplest form of division of labour. The argument remains basically unchanged if we increase the number of cells in the above model.

Two aspects of the model

Our model as expressed by eqs.(2) consists of two parts of different nature: one representing continuous changes of the variables and the other involving discrete changes of the states. When we are interested in the *process* leading to cell differentiation, we may consider only the dynamics of physiological state during the period before the gene state changes. If we can further postulate that the cell interaction is mediated by metabolites diffusing in the tissue, or by a process that can be described by a diffusion equation, the model becomes a reaction-diffusion type.

On the contrary, in the cases, such as *tissue proportioning* and *pattern formation*, which involve more than one cell-types, we will be more interested in the changes of gene state, rather than the dynamics of physiological parameters, for, in such cases, the stability of the coexistence of different cells-types will be of primary importance. Since reaction-diffusion systems have been a subject of extensive investigations, we concentrate hereafter on the latter aspect of the model.

Factors influencing the expression of new genes

To be more specific, consider a hypothetical cell with two physiological variables, p and q, and suppose there are N such cells in the system. As pointed out earlier, not all the components of the physiological state will directly contribute to cell interaction. Here we assume that q is the component of the physiological state that directly participates in cell interaction, while p represents the component involved in the cell-autonomous dynamics. The former components will hereafter be called "intercellular signals". Assuming that all the cells of the system express gene C in the beginning, we ask how

many of them come to express the gene of interest, gene A, in addition to gene C. For the convenience's sake, the cell expressing only C will hereafter be called C-cell and the cell expressing both C and A, A-cell.

Suppose one of the cells, cell *1*, is about to express gene A. There are three factors that influence the expression of gene A in cell *1*;

(1) its own physiological state,

(2) physiological state of other C-cells,

(3) physiological state of A-cells if they exist in the system.

Each of these factors has either activating, inhibiting, or no influence on the expression of gene A of cell *1*, and whether it is switched on or not will be determined by the sum of the effects of the factors (1) - (3).

Control of cell-type proportions

Consider a system comprising N cells in which cell dynamics and cell interaction are described by single parameters p and q, respectively. Then the model can be written as

$$G_i(t) = \begin{cases} \{C\} & \text{if} \quad p_i < p^* \\ \{C, A\} & \text{if} \quad p_i \geq p^* \end{cases} \tag{3a}$$

$$\frac{dp_i(t)}{dt} = v(p_i(t);\, G_i(t),\, Q_i) \tag{3b}$$

where

$$Q_i = \sum_{j=1}^{N} r_{ij} q_j,$$

$$q_j = \begin{cases} c & \text{if} \quad G_j = \{C\} \\ a & \text{if} \quad G_j = \{C, A\}. \end{cases}$$

Here, r_{ij} represents the efficiency of the transmission of the jth cell's effect (q_j) to the ith cell. The initial conditions are $p_i(0) = p_i^0 \, (< p^*)$, $G_i(0) = \{C\}$. Now we examine the conditions for division of labour. Although we deal with

discrete phenomena (ON $\leftrightarrow$ OFF switches), we cannot totally ignore the physiological change, since any change of the gene state must be preceded by a change in the physiological state. If $\exists p_i < p^*$ which satisfies

$$v(p_i(t);\ \{C\},\ c \cdot \sum r_{ij}\ ) \leq 0$$

for all i's, all cells remain to be C-cells. If

$$v(p^*;\ \{C\},\ a \cdot \sum r_{ij}\ ) > 0,$$

holds for all i's, then all the cells become A-cells. If otherwise, division of labour can result. The number of A-cells, N^A, is calculated from

$$v(p^*,\ \{C\},\ c \cdot \sum_{C\text{-cells}} r_{ij} + a \cdot \sum_{A\text{-cells}} r_{ij}) = 0. \tag{4}$$

In the simple case where $r_{ij} = r$ holds for all (i, j), $Q = (N^C c + N^A a)r$, and we have division of labour if

$$a \leq 0 < c \tag{5a}$$

and

$$\begin{aligned} &Q^*/rc < N && \text{if } Q^* \geq 0 \\ &Q^*/ra \leq N && \text{if } Q^* < 0 \end{aligned} \tag{5b}$$

hold, where Q^* satisfies $v(p^*, \{C\}, Q^*) = 0$. Here, without loss of generality $\partial v/\partial Q > c$ was assumed. Inequality (5a) indicates that for division of labour to occur the intercellular signals given off by C-cells need to promote the expression of gene A whereas that of A-cells must be inhibitory to it, whereas inequalities (5b) shows the presence of a lower limit of the number of cells for division of labour.to occur (Fig.3).

The proportion of A cells is calculated to be

$$\frac{N^A}{N} = \frac{c}{c - a} - \frac{Q^*}{r(c - a)} \cdot \frac{1}{N}, \tag{6}$$

which converges to a constant value $c/(c - a)$ for large N, *i.e.* constancy of cell-type proportions holds if the system is sufficiently large.

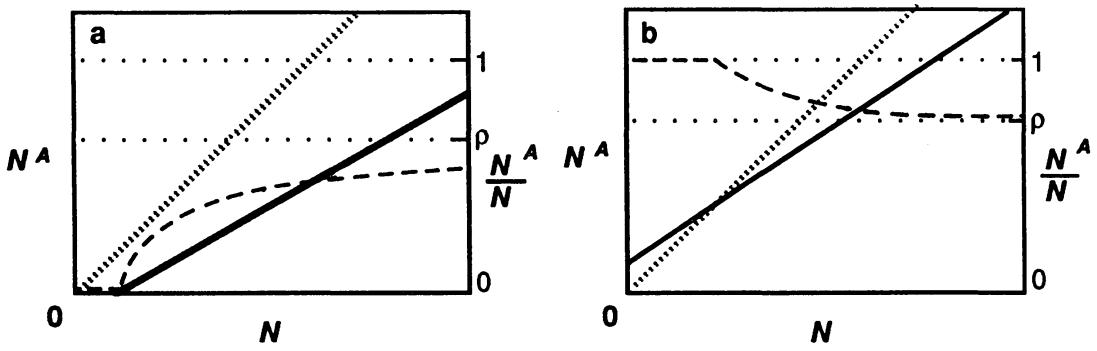

Fig.3. Schematic graphs showing the dependency of the cell interaction parameter (Q) on the total cell numbers (N) and the number of A-cells (N^A). Proportion of A-cells is also shown as a function of N. **a**, $Q^* > 0$; **b**, $Q^* \leq 0$. $Q = Nc + N^A(a - c)\}r$. Solid line, $Q = Q^*$; dotted line, $N = N^A$; dashed line, proportion. $\rho = c/(c - a)$.

In real developmental systems, the cells, even when they have the same gene state, are in general not identical to each other, and the time courses of their physiological changes will also be non-identical, so that we can conceive that some cells differentiate into A-cells earlier than others. Such a difference results from differences in the dynamics governing the physiological change. If such a heterogeneity in v is taken into account, the proportion is obtained from the distribution of Q^* within the cell population, $N = F(Q^*)$, and eq.(6).

The above argument postulates an equal efficiency of the transmission of the intercellular signal q irrespective of the cell state (*i.e.* same r for c and a). In real developmental systems, there are cases in which one or more new intercellular signals come into play upon the expression of new genes. By way of example, suppose a system consisting of N cells in which A-cells give off a new intercellular signal, in addition to c, as a result of the expression of gene A. The efficiency of transmission of this signal, a , will in general be different from that for c. We designate these by r^A and r^C, respectively. These signals will act on different reactions in the cell dynamics. The interaction parameter Q is therefore separated into $\{Q^C, Q^A\}$. For clarity,we consider the case where v depends on Q's linearly. Then

$$v(p, \{C\}, Q^C, Q^A) = v(p, \{C\}, 0, 0) + s^C Q^C + s^A Q^A$$
$$= v^0 + N s^C r^C c + N^A s^A r^A a, \qquad (7)$$

where s^C and s^A are the sensitivities of the cell to the effects c and a, respectively. The conditions for division of labour can be derived in a similar manner as described above:

$$a + c \leq 0 < c$$

and (8)

$$Ns^C r^C c + N^A s^A r^A a \leq -v^0 < Ns^C r^C c.$$

By equating v to 0 for $p = p^*$, we obtain

$$\frac{N^A}{N} = \frac{s^C r^C c}{-s^A r^A a} + \frac{v^0(p^*)}{-s^A r^A a} \cdot \frac{1}{N} \cdot \tag{9}$$

In eq.(9), it can be seen that constant proportion holds for large N under the condition (8). Six factors are identified which influence the proportion: the effects of gene C and gene A (c, a), the efficiency of transmission of these effects (r^C , r^A), and the sensitivities of the cells to these effects (s^C , s^A). For instance, the larger the inhibition by A-cells of other cells' expression of gene A, the lower the proportion of A-cells.

Stability and proportion regulation

In the above examples, expression of gene A, and therefore division of labour also, are stable if

$$v(p, \{A\}, Q^*) > 0 \qquad \text{for } p \geq p^* \tag{10}$$

holds. The proportion of A-cells is regulated automatically. If, for example, all or part of of the A-cells are removed from the system, the average level of a, which has been suppressing the emergence of excessive A-cells, becomes lower than at the equilibrium (*i.e.* $Q > Q^*$), and consequently part of the C-cells come to express gene A so that the proportion of A-cells would be restored. On the other hand, removal of C-cells may not induce regulation. Removal of C-cells causes Q^* to decrease. However, for A-cells to dedifferentiate (*i.e.* to switch off gene A) to regenerate C-cells, Q^* needs to become sufficiently small so that $Q^* < Q^{*'}$ ($Q^{*'}$ is defined by $v(p^* \{A\}, Q^{*'}) = 0$).

Hence $v(p, \{A\}, N^A ar) > 0$ is required for regulation to occur after removal of C-cells. It follows from $Q^{*\prime} > Q^*$ and $\partial v/\partial N^A < 0$ that the new proportion of A-cells after the regulation induced by removal of C-cells is generally smaller than the initial proportion.

Discussion

We have concentrated in the preceding arguments on the problems of cell-type proportion. Formation of spatial patterns by differentiated cells is another important aspects of multicellular development. Most existing mathematical models aim at producing non-uniform distributions of the "morphogen" in a continuous field. There are cases in which a specific spatial pattern arises within the continuum of cytoplasm, such as in the early development of *Drosophila*, which will be described by a set of equations, defined on a continuous field, that represent the chemical reactions and diffusion of the molecules involved. In multicellular organisms, on the other hand, a pattern is formed by discrete units (cells) each of which taking, roughly speaking, one state from a set of discrete states. To deal with the problems of multicellular development such as cell-type proportioning and pattern formation, there is no reason, therefore, for adhering to dynamical systems on continuous space such as ordinary reaction-diffusion systems. The present model, on the other hand, is based on discrete units, and, by placing some additional constraints on r_{ij}, it proves to be useful in studying pattern formation. For instance, by assuming that r_{ij} is reversibly proportional to the square of distance, the model can be seen as modelling a tissue structure in which cell interaction is mediated by diffusible substances (for reviews on diffusible morphogens, see *e.g.* Kay & Smith, 1989). With such a model, it can be shown that the widely-accepted principle of short-ranging activation and long-ranging inhibition (Meinhardt, 1982) is not the universal feature of the systems showing a stable coherent pattern.

The unit of the system has been called the "cell" throughout this paper, implicating that the model is specifically concerned with cell differentiation. The present model, however, may be applied to a variety of biological systems in which "division of labour" arises. What we called the "cell" may be the actual cell, a group of cells which behaves as a well defined unit (such as a segment of the arthropod) or an individual in the society (such as an individual in social insects). The applicability of the model will be further extended by generalizing its formulation in appropriate ways.

References

Kay, R. & Smith, J. eds. (1989). The molecular basis of positional signalling. *Development*, 1989 Supplement.

Meinhardt, H. (1982). "Models of Biological Pattern Formation." Academic Press, London.

Murray, J. D. (1989). "Mathematical Biology." Springer-Verlag, Berlin.

Nagorcka, B. N. (1989). Wavelike isomorphic prepatterns in development. *J. theor. Biol.* **137**, 127-162.